# Université de Genève

## Département d'informatique

**Théorie de l'information pour la science des données et l'apprentissage automatique**

Série d'exercices 5 — 29.03.2023 Délai : 04.04.2023

---

**Note :** Veuillez suivre les instructions suivantes
`https://moodle.unige.ch/mod/page/view.php?id=171850`.

**Entropie** L'incertitude ou entropie d'une variable aléatoire discrète (RV en anglais) $X$ qui prend des valeurs dans l'ensemble $\mathcal{X}$ (aussi appelé alphabet $\mathcal{X}$) est définie par :

$$H(X) \triangleq \mathbb{E}_{P_X}\left[-\log_2 P_X(X)\right] = -\sum_{x\in\mathcal{X}} P_X(x)\log_2 P_X(x),$$

où $P_X(\cdot)$ désigne la fonction de masse (PMF en anglais) de $X$.

**Entropie Conjointe** L'incertitude ou entropie conjointe d'un vecteur aléatoire discret $(X,Y)^T$ est définie par :

$$H(X,Y) \triangleq \mathbb{E}_{P_{X,Y}}\left[-\log_2 P_{X,Y}(X,Y)\right] = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{X,Y}(x,y)\log_2 P_{X,Y}(x,y),$$

où $P_{X,Y}(\cdot,\cdot)$ désigne la fonction de masse conjointe de $(X,Y)$.

**Entropie Conditionnelle** L'incertitude ou entropie conditionnelle d'une RV $X$ étant donné une RV $Y$ est définie par :

$$\begin{aligned} H(X \mid Y) &\triangleq \mathbb{E}_{P_{X,Y}}\left[-\log_2 P_{X|Y}(X \mid Y)\right] \\ &= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{X,Y}(x,y)\log_2 P_{X|Y}(x \mid y). \end{aligned}$$

**Règle de la chaîne pour l'Entropie** Soit $n$ RV discrètes $X_1, ..., X_n$ avec une PMF conjointe $P_{X_1,...,X_n}$. Alors

$$\begin{aligned} H(X_1, X_2, \cdots, X_n) &= H(X_1) + H(X_2 \mid X_1) + \cdots + H(X_n \mid X_1, X_2, \cdots, X_{n-1}) \\ &= \sum_{k=1}^{n} H(X_k \mid X_1, X_2, \cdots, X_{k-1}) \end{aligned}$$

La règle de la chaîne ci-dessus suit directement de la règle de la chaîne pour les PMFs :

$$P_{X_1,X_2,\cdots,X_n} = P_{X_1}\, P_{X_2|X_1}\, P_{X_3|X_1,X_2} \cdots P_{X_n|X_1,...,X_{n-1}}.$$

**Problème 1 (40 points).** Soit deux variables aléatoires binaires $X$ et $Y$ avec une probabilité conjointe donnée par la table ci-dessous :

| X | Y | $p_{XY}(x,y)$ |
|---|---|---|
| 0 | 0 | $\frac{1}{4}$ |
| 0 | 1 | 0 |
| 1 | 0 | $\frac{1}{4}$ |
| 1 | 1 | $\frac{1}{2}$ |

Calculez tout d'abord les probabilités marginales et conditionnelles. Puis calculez :

(a) $H(X)$

(b) $H(Y)$

(c) $H(X|Y)$

(d) $H(Y|X)$

(e) $H(X,Y)$

(f) $H(X,Y) - H(X)$

(g) $H(X,Y) - H(Y)$

**Problème 2 (20 points).** Imaginons quatre variables aléatoires discrètes $X_1, X_2, X_3, X_4$ avec les relations statistiques telles qu'indiquées par les deux diagrammes ci-dessous. Pour chaque cas, obtenez la forme la plus simple de :

i. $p(x_1, x_2, x_3, x_4)$

ii. $H(X_1, X_2, X_3, X_4)$

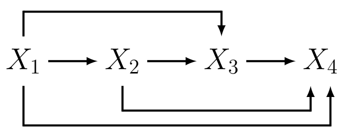


Figure 1:

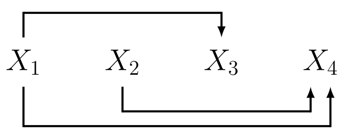


Figure 2:

**Problème 3 (20 points).** Soit deux variables aléatoires $X$ et $Y$. Montrez que :

$$H(X|Y) \leqslant H(X).$$

**Problème 4 (20 points).** Soient des variables aléatoires arbitraires $X_1 \cdots X_n$. Montrez que :

$$H(X_1, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i).$$